

Financial LPPL Bubbles with Mean-Reverting Noise in the Frequency Domain

Vincenzo Liberatore *

February 1, 2011

Abstract

The log-periodic power law (LPPL) is a model of asset prices during endogenous bubbles. A major open issue is to verify the presence of LPPL in price sequences and to estimate the LPPL parameters. Estimation is complicated by the fact that daily LPPL returns are typically orders of magnitude smaller than measured price returns, suggesting that noise obscures the underlying LPPL dynamics. However, if noise is mean-reverting, it would quickly cancel out over subsequent measurements. In this paper, we attempt to reject mean-reverting noise from price sequences by exploiting frequency-domain properties of LPPL and of mean reversion. First, we calculate the spectrum of mean-reverting Ornstein-Uhlenbeck noise and devise estimators for the noise's parameters. Then, we derive the LPPL spectrum by breaking it down into its two main characteristics of power law and of log-periodicity. We compare price spectra with noise spectra during historical bubbles. In general, noise was strong also at low frequencies and, even if LPPL underlied price dynamics, LPPL would be obscured by noise.

1 Introduction

Financial bubbles and busts have devastating effect on the economy and on markets. However, the existence and characteristics of bubbles are notoriously hard to ascertain if not with hindsight. This paper contributes to the investigation of financial bubbles within the LPPL framework [Sornette, 2004], and specifically it examines its frequency-domain properties under a mean-reverting noise model.

Endogenous financial bubbles have been modeled as a *log-periodic power law (LPPL)* [Sornette, 2004]. The LPPL model has two main characteristics:

- Super-exponential growth, leading to a *critical time* at which the asset price will burst (*power law*), and
- Oscillations that become progressively faster as the critical time approaches (*log-periodicity*).

Super-exponential growth is a sign that price growth is unsustainable. The oscillatory behavior indicates an incipient system failure, and is often associated with increasingly more rapid and pronounced swings in investor sentiment [Stauffer and Sornette, 1998].

*Division of Computer Science, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7071, USA. E-mail: vl@case.edu. URL: <http://vincenzo.liberatore.org/>.

Previous work has paid considerable attention to fitting the LPPL law to historical time series of financial bubbles, and a recent summary reviews the state of the art [Jiang et al., 2010]. Current methods for LPPL bubble detection have been tested in an on-line experiment [Sornette et al., 2009]. Previous work mostly focus on the two related issues of statistical significance and of noise in the data. The LPPL model parameters can be fit to the price sequences via non-linear least-squares [Jiang et al., 2010], and we have previously proposed efficient parallel algorithms for least-square fitting [Liberatore, 2010]. The least-square algorithm returns an estimate for the LPPL parameters as well as a residual error. However, least squares always estimate parameters for the LPPL model, regardless of whether LPPL underlies actual price dynamics or not. The least square residual error gives a sense of the presence of LPPL: if the final mean squared error is significantly smaller than the error at the beginning of the numerical fit, then there is some evidence for an underlying LPPL process. However, reliance on mean squared errors can produce spuriously high goodness of fit [Granger and Newbold, 1974, Phillips, 1986]. Although LPPL is often visible intuitively in price series and non-linear least squares substantially reduce the mean squared error over an initial exponential fit, the log-periodic component of the S&P 500 is not statistically significant prior to the 1987 crash if the last year of data is removed [Feigenbaum, 2001]. Furthermore, Bayesian methods fundamentally reject the hypothesis that LPPL underlies price dynamics during a bubble [Chang and Feigenbaum, 2006]. An alternative approach is to verify the non-linear least squares using unrelated tests, such as bounding the range of acceptable LPPL parameters or using a Lomb transform [Jiang et al., 2010]. However, log-periodic variations prior to large drawdowns fail to satisfy the parameter restrictions in the LPPL model [Chang and Feigenbaum, 2008].

In fitting LPPL to prices, the fundamental problem is that the daily LPPL returns are typically orders of magnitude smaller than the measured price returns. In other words, a hypothetical LPPL signal would be swamped by overwhelming noise, and thus it is difficult to ascertain its presence and its parameter values. Furthermore, it is also possible that the noise variance increases over time, in which case the observed prices tend to progressively differ more significantly from an underlying deterministic LPPL. When noise variance increases, the noise effectively becomes the signal. As such, LPPL estimation is plagued by a low signal-to-noise ratio, which additionally may be decreasing over time. However, an LPPL model with mean-reverting noise (e.g., Ornstein-Uhlenbeck) has been recently proposed [Lin et al., 2009]. If noise is mean-reverting, it has bounded variance, and thus the measured prices fits more tightly around the underlying LPPL. Furthermore, although a mean-reverting process may be characterized by large daily return, the daily gyrations will rapidly cancel each other out. In other words, mean-reverting noise is characterized by relatively low power at low frequencies. Thus, mean-reversion would make it possible to reconstruct the underlying LPPL by focusing on the lower frequency components of measured prices.

In this paper, we examine LPPL with mean-reverting Ornstein-Uhlenbeck noise in the frequency domain. The paper investigates the mean-reverting properties of the noise, and whether the de-noised power spectrum shows a deterministic LPPL signature. Section 2 describes the LPPL model. Section 3 gives the necessary background on Ornstein-Uhlenbeck processes, analyzes its spectrum, and discusses two methods for estimating its parameters. Section 4 gives our methodology for calculating spectra of finite price sequences. Section 5 discusses the LPPL spectrum by breaking it down into its two main features of power law and log-periodicity. Section 6 discusses noise rejection in noisy LPPL price sequences. Section 7 evaluates the approach on major historical bubbles. Section 8 concludes the paper and outlines possible future work.

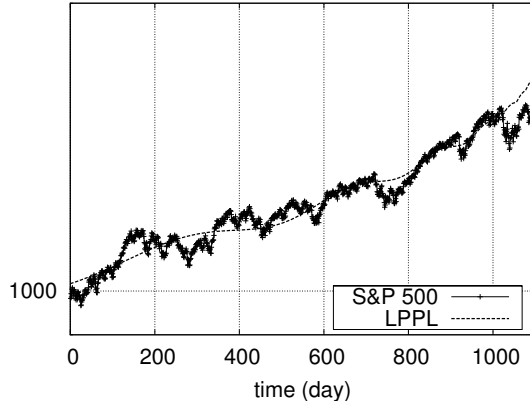


Figure 1: Log-scale S&P 500 and an LPPL fit from July 2003 to June 2007.

2 LPPL

The *log-periodic power law* is a function:

$$\ell(t) = A - B(T - t)^m(1 + C \cos(\omega \ln(T - t) + \phi)) , \quad (1)$$

where $B > 0$ and $0 < m \leq 1$. The LPPL function is a model for a sequence of price logarithms $p(0), p(1), \dots, p(n-1)$ in the sense that $p(i) = \ell(i) + \nu(i)$, where $\nu(i)$ is random noise independent of ℓ . If we assume that $\ell(T) = A$, the LPPL function ℓ is defined for $t \leq T$. In particular, since ℓ is a model for a finite sequence of prices, we will also restrict ℓ over $t \geq 0$. As matter of notational convention, we remark that we call price sequence and denote as $p(0), p(1), \dots, p(n-1)$ what in reality is the natural logarithm of prices. The convention makes our notation more compact, but it should be kept in mind that, for example, our daily price returns are basically price ratios. Figure 1 shows the S&P 500 daily closing prices and a least squares LPPL fit for the four years between July 2003 and June 2007. If $m = 1, C = 0$, LPPL reduces to an exponential model of the price sequence. If $m < 1$, then LPPL grows super-exponentially until the *critical time* T . If $C, \omega > 0$, then LPPL exhibits oscillations that become progressively more frequent as t approaches the critical time. Since ℓ is a deterministic signal and ν is a wide-sense stationary process independent of ℓ , most properties of stationary signals can be applied [Pollock, 2001]. Since ν is independent of ℓ , $|P(f)|^2 = |L(f)|^2 + S(f)$, where $L(f)$ is the transform of ℓ , $P(f)$ is the transform of p , and $S(f)$ is the noise power.

3 Ornstein-Uhlenbeck Noise

3.1 Background

The *Ornstein-Uhlenbeck process* $\nu(t)$ is the solution of the stochastic differential equation:

$$\nu'(t) = -\frac{\nu(t)}{\tau} + \sigma \Gamma(t) ,$$

where $\Gamma(t)$ is Gaussian white noise, $\tau \geq 0$ is called the *relaxation time*, and $\sigma \geq 0$ is called the *diffusion constant* [Bibbona et al., 2008]. The Ornstein-Uhlenbeck process can be interpreted as white noise that has been filtered by a first-order system with cut-off frequency $f_c = 1/(2\pi\tau)$. When $\tau \rightarrow \infty$, the Ornstein-Uhlenbeck process reduces to a Wiener process. As $t \rightarrow \infty$, $\text{Var}[\nu(t)] = \sigma^2\tau/2$ and $\text{Cov}[\nu(t+h), \nu(t)] = \sigma^2\tau e^{-|h|/\tau}/2$ [Bibbona et al., 2008]. A Ornstein-Uhlenbeck process can be simulated by means of:

$$\nu(t+1) = \nu(t)e^{-1/\tau} + \sqrt{\frac{\sigma^2\tau}{2}(1 - e^{-2/\tau})}u_t, \quad (2)$$

where u_t is a sample value of a normal random variable [Gillespie, 1991]. The stationary Ornstein-Uhlenbeck process can be simulated by using recurrence (2) starting from $\nu(0)$ being the sample of a Gaussian random variable with variance equal to $\lim_{t \rightarrow \infty} \text{Var}[\nu(t)] = \sigma^2\tau/2$.

3.2 Spectrum

The two-sided power spectrum is defined for $t \rightarrow \infty$, is equal to [Wang and Uhlenbeck, 1945]:

$$T(f) = \frac{\sigma^2\tau^2}{1 + 4\pi^2\tau^2 f^2} \quad \left(0 \leq f \leq \frac{1}{2}\right)$$

and it is the same as that of a first-order low-pass filter with cut-off frequency f_c . Since we will be dealing with sampled Ornstein-Uhlenbeck processes, it is also useful to derive the spectrum of the discrete-time Fourier transform. The transform of the auto-correlation function is

$$\begin{aligned} S(\omega) &= \sum_{h=-\infty}^{\infty} \frac{\sigma^2\tau}{2} e^{-|h|/\tau} e^{-j\omega h} \\ &= \frac{\sigma^2\tau}{2} \left(-1 + \sum_{h=0}^{\infty} e^{-h/\tau} (e^{-j\omega h} + e^{j\omega h}) \right) \\ &= -\frac{\sigma^2\tau}{2} + \sigma^2\tau \sum_{h=0}^{\infty} a^h \cos(\omega h), \end{aligned}$$

where $a = e^{-1/\tau}$. Since, it is known that for $|r| < 1$ [Råde and Westergren, 1999, eq. 13.2(29)]

$$\sum_{h=0}^{\infty} r^h \cos(hk) = \frac{1 - r \cos k}{1 - 2r \cos k + r^2},$$

we have that

$$S(\omega) = \frac{\sigma^2\tau}{2} \left(-1 + 2 \frac{1 - a \cos \omega}{1 - 2a \cos \omega + a^2} \right) = \frac{\sigma^2\tau}{2} \frac{1 - a^2}{1 - 2a \cos \omega + a^2}.$$

In terms of frequency $f = \omega/(2\pi)$,

$$S(f) = \frac{\sigma^2\tau}{2} \frac{1 - a^2}{1 - 2a \cos(2\pi f) + a^2}.$$

Figure 2 shows the power of the discrete-time Fourier transform $S(f)$ and of the continuous Fourier transform $T(f)$ for $\tau = 5$ and $\sigma = 1/5$. The power spectra are similar except at the highest frequencies. At low frequencies, since $a \simeq 1 - 1/\tau$ and $a^2 \simeq 1 - 2/\tau$, $T(f) \simeq S(f) \simeq \sigma^2/(1 - a)^2 \simeq \sigma^2\tau^2$.

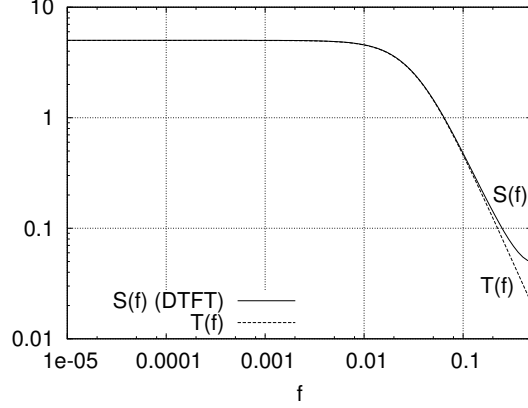


Figure 2: Power spectrum of the Fourier and discrete-time Fourier transform for $\tau = 5$, $\sigma = 1/5$.

3.3 Ornstein-Uhlenbeck Estimation

We now turn to the estimation of the parameters σ and τ in the Ornstein-Uhlenbeck noise. We will follow two approaches for parameter estimation. The first approach is based on the intuition that on short time frames the noise dominates the LPPL signal (e.g., [Chang and Feigenbaum, 2006]). Therefore, the noise parameters could be determined with high accuracy by looking at diffusion and mean reversal over short intervals, possibly corrected by a first-order trend estimate. The second approach is pessimistic: it assumes $\tau \rightarrow \infty$ and finds the largest value of σ that is consistent with the observed signal. In other words, the pessimistic approach attempts to explain the observed signal as much as possible as Wiener noise.

3.3.1 Maximum Likelihood

High-frequency estimation is based on the assumption that over short time scales the signal is completely dominated by noise. In this case, $\ell(t+1) \simeq \ell(t)$, and since $p(t) = \ell(t) + \nu(t)$, we have that $\nu(t+1) - \nu(t) \simeq p(t+1) - p(t)$.

Given a sequence of observations $\nu(0), \nu(1), \dots, \nu(n-1)$ of an Ornstein-Uhlenbeck process, the maximum likelihood estimates of τ and σ^2 are

$$\hat{\tau} = \ln \frac{S_{xx}}{S_{xy}} ,$$

$$\hat{\sigma}^2 = \frac{2\hat{\tau}}{1 - \hat{a}^2} \frac{1}{n-1} (S_{yy} - 2\hat{a}S_{xy} + \hat{a}^2S_{xx}) ,$$

where $\hat{a} = e^{-1/\hat{\tau}}$, $S_{xx} = \sum_{i=1}^{n-1} \nu^2(i-1)$, $S_{xy} = \sum_{i=1}^{n-1} \nu(i-1)\nu(i)$, and $S_{yy} = \sum_{i=1}^{n-1} \nu^2(i)$ [van den Berg, 2007]. Maximum likelihood estimation is an estimation technique that operates exclusively in the time-domain.

However, the ν values are only known indirectly through the differences $\nu(t+1) - \nu(t)$. Define $\alpha_t = \nu(t) - \nu(0)$ and observe that α_t is known via the telescoping summation $\nu(t) = \nu(0) + \sum_{i=0}^{t-1} (\nu(i+1) - \nu(i))$. Since the Ornstein-Uhlenbeck process has no drift, $S_x S_{xy} = S_y S_{xx}$, where

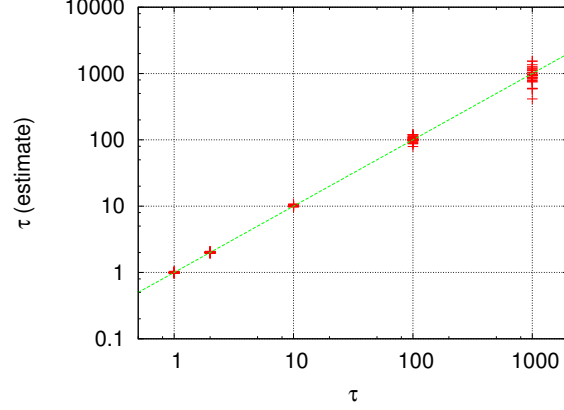


Figure 3: Estimate $\hat{\tau}$ versus its underlying value τ . For each value of τ , 30 Monte Carlo simulations were executed over intervals of length $n = 25000$.

$S_x = \sum_{i=1}^{n-1} \nu(i-1)$ and $S_y = \sum_{i=1}^{n-1} \nu(i)$ [van den Berg, 2007]. Therefore,

$$\begin{aligned}
0 &= \sum_{i=1}^{n-1} (\nu(0) + \alpha_{i-1}) \sum_{j=1}^{n-1} (\nu(0) + \alpha_{j-1})(\nu(0) + \alpha_j) - \sum_{i=1}^{n-1} (\nu(0) + \alpha_i) \sum_{j=1}^{n-1} (\nu(0) + \alpha_{j-1})^2 \\
&= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} (\nu(0)^3 + \nu(0)^2(\alpha_j + \alpha_{j-1} + \alpha_{i-1}) + \nu(0)(\alpha_j\alpha_{j-1} + \alpha_{i-1}\alpha_j + \alpha_{i-1}\alpha_{j-1}) \\
&\quad + \alpha_{i-1}\alpha_j\alpha_{j-1}) \\
&\quad - \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \nu(0)^3 + \nu(0)^2(2\alpha_{j-1} + \alpha_i) + \nu(0)(\alpha_{j-1}^2 + 2\alpha_i\alpha_{j-1}) + \alpha_i\alpha_{j-1}^2 \\
&= (n-1)\nu(0)A_{xy} + \nu(0)A_x^2 + A_xA_{xy} - (n-1)\nu(0)A_{xx} - \nu(0)A_xA_y - A_yA_{xx},
\end{aligned}$$

where $A_x = \sum_{i=1}^{n-1} \alpha_{i-1}$, $A_y = \sum_{i=1}^{n-1} \alpha_i$, $A_{xy} = \sum_{i=1}^{n-1} \alpha_i\alpha_{i-1}$, and $A_{yy} = \sum_{i=1}^{n-1} \alpha_i\alpha_i$, Solving for $\nu(0)$ now gives

$$\nu(0) = \frac{A_yA_{xx} - A_xA_{xy}}{(n-1)A_{xy} - (n-1)A_{xx} + A_x^2 - A_yA_x}.$$

Given the estimate for $\nu(0)$ and the α_i 's, the $\nu(i)$'s can be estimated, and thus τ and σ^2 .

To increase the estimate accuracy in the presence of an underlying deterministic trend ℓ , we estimate a linear fit to the log-prices p and subtract it from $p(t)$.

Monte Carlo simulations were conducted to ascertain the accuracy of the maximum likelihood estimation. The estimation was quite accurate when the underlying LPPL trend gave negligible contributions to the α_i 's, but less accurate otherwise. Additionally, even in the complete absence of an underlying trend, the estimate $\hat{\tau}$ of τ became progressively less accurate as τ gets larger (Figure 3). An intuitive explanation is that when τ is large, then mean reversion affects only slightly the α_i 's, making it hard to obtain an accurate fix on τ . In principle, the problem could be obviated by formulating an estimate with returns over longer periods of time (i.e., $\tilde{\nu}(i) = \nu(ki)$, $k > 1$). However, the problem is compounded if an underlying trend ℓ has relatively large power at frequencies smaller

than the critical frequency $f_c = 1/(2\pi\tau)$. In this case, the underlying trend ℓ affects the signal at the same frequencies at which the relaxation term does, thus obscuring the exact contribution of the relaxation constant. The case when τ is large will be addressed by pessimistic estimation.

3.4 Pessimistic Estimation

In pessimistic estimation, we assume that the noise is generated by a Weiner process (i.e., $\tau \rightarrow \infty$) and estimate σ using the constraint $|P|^2 = |L|^2 + S \geq S$ that the noise power should not exceed the observed signal power. The estimate is pessimistic in that it assumes no mean reversion and thus strong noise even at low frequencies.

Let $S_\tau(f)$ be the power spectrum of an Ornstein-Uhlenbeck process when the relaxation constant is τ and define $S_\infty(f) = \lim_{\tau \rightarrow \infty} S_\tau(f)$. Then, if $f > 0$,

$$S_\infty(f) = \lim_{\tau \rightarrow \infty} S_\tau(f) = \frac{\sigma^2}{4(1 - \cos(2\pi f))} \lim_{\tau \rightarrow \infty} \tau (1 - e^{-2/\tau}) = \frac{\sigma^2}{2(1 - \cos(2\pi f))} ,$$

and

$$S_\infty(0) = \lim_{\tau \rightarrow \infty} \frac{\sigma^2 \tau}{2} \frac{1 - a^2}{(1 - a)^2} = \sigma^2 .$$

It should be the case that $P(f) \leq S_\infty(f)$ for all f 's. In practice, however, the observed signal is the realization of a stochastic signal, and so $P(f)$ may or may not exceed $S_\infty(f)$ [Loève, 1978]. Hence, we only impose integral constraints of the form

$$\int_{f^-}^{f^+} S_\infty(f) df \leq \int_{f^-}^{f^+} P(f) df ,$$

where f^- and f^+ are parameters to be determined with $0 < f^- < f^+ \leq 1/2$. Furthermore, the spectrum P is calculated only at the discrete points $f_i = i/n$, so the constraint becomes

$$\sum_{i=h}^k S_\infty(f_i) \leq \sum_{i=h}^k P(f_i) ,$$

which then leads to

$$\sigma^2 \leq \frac{2 \sum_{i=h}^k P(f_i)}{\sum_{i=h}^k 1/(1 - \cos(2\pi f_i))} . \quad (3)$$

A pessimistic estimate can be obtained by calculating the bounds for $h_0 = 1$, $k_j = h_{j+1} = (1 + \alpha)h_j$ (for some constant $\alpha > 0$, e.g., $\alpha = 1$) and then taking the most conservative bound.

The numerical calculation of the bounds is complicated by the fact that the summations in (3) involve terms whose value can differ by orders of magnitude. In this case, if the largest terms are added first, due to the finite representation precision, the smallest terms are likely to be truncated. In an extreme scenario, many small terms could collectively dominate the value of the summation but since they are added individually to a larger partial sum, their contribution would be lost. In general, a summation is numerically more accurate when it always adds terms of comparable magnitude. Since S_∞ is a decreasing function (and the same holds in the expectation for P), the numerical calculation should always start from the largest frequency f_k and proceed backward to f_h . (As an aside, however, the numerical calculation is only needed for the numerator, since the denominator could be replaced by an integral that is solvable analytically.)

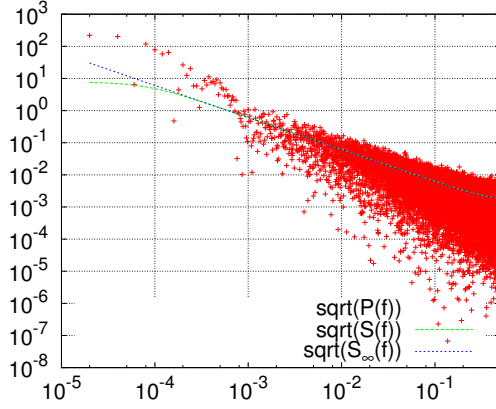


Figure 4: DFT amplitude of a log-periodic function, ($A = 10, B = 0.008, T = 26000, m = 0.7, n = 25000, C = 0.4, \omega = 2\pi, \phi = \pi$), underlying Ornstein-Uhlenbeck noise ($\sigma^2 = .000015, \tau = 2000, f_c = 7.96 \cdot 10^{-5}$), and pessimistic estimate ($\alpha = 1, \sigma^2 = .0000145$).

Figure 4 shows an example of pessimistic estimate in a case when $\tau = 2000$ is relatively large and difficult to estimate because $P(f_c)$ is dominated by the power of the underlying signal ℓ . The pessimistic estimate was quite accurate, and S_∞ differs from the underlying $S(f)$ only for the smallest frequency values. Similar conclusions were supported by several additional Monte Carlo simulations.

4 Reflection

The Ornstein-Uhlenbeck process describes an infinitely long signal. On the contrary, $p(0), p(1), \dots, p(n-1)$ is a finite sequence. Similarly, ℓ is defined in the finite interval $[0, T]$. In the calculation of the spectrum, an ambiguity arises regarding the behavior of p and ℓ outside of their definition interval. In many applications, finite sequences are multiplied by a window function, or they are implicitly assumed to be periodic. Implicit periodization matches directly the discrete Fourier transform (DFT), in the sense that, given the sequence $p(0), p(1), \dots, p(n-1)$, the DFT assumes that the signal is infinite, periodic, with period n , and that $p(i) = p(i-n)$ for $i = 0, \pm 1, \pm 2, \dots$. In typical log-periodic sequences, $p(n-1) \gg p(0)$, and if the signal were viewed as periodic, then there would be a large step between $p(n-1)$ and $p(n) = p(0)$. The step is primarily an artifact of the implicit DFT assumption of periodicity, but the large artificial step could dominate the behavior of the power spectrum. As for windowing, typical window functions either leave a large boundary step (e.g., rectangular window) or disregard the fact that the LPPL behavior becomes more pronounced at the boundaries of the observation window (e.g., Hann).

In this paper, the discontinuity is obviated by *reflection*: given a sequence $p(0), p(1), \dots, p(n-1)$, a new sequence is constructed by juxtaposing the original sequence and a reverse version of the same sequence: $p(n-1), \dots, p(1), p(0), p(1), \dots, p(n-2)$, and assuming that the new sequence is periodic with period $2(n-1)$. The reflected sequence has no abrupt discontinuity between the first and last element, and thus no such step is visible in its spectrum. We also tried alternative approaches, such as appending after $p(n-1)$ values that smoothly interpolate between $p(n-1)$ and

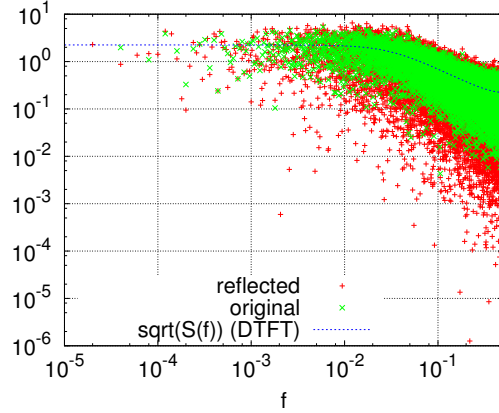


Figure 5: DFT amplitude of a realization ($\sigma = .2$, $\tau = 5$) of the Ornstein-Uhlenbeck process, of its reflection, and $\sqrt{S(f)}$.

$p(0)$ over a long time frame, but we did not find significant differences between these alternatives and reflection.

To reiterate, given a finite sequence or a function defined over a finite interval, the Fourier transform always requires an assumption of some sort on the sequence behavior outside the given interval. For example, if the sequence is left unchanged, the DFT implicitly assumes that the sequence is periodic. In this light, reflection is a natural choice because, unlike implicit periodization or windowing, it eliminates abrupt discontinuities at the boundaries. Reflection also lead to a symmetrical analysis of bubbles and anti-bubbles (i.e., a bubble-like behavior in reverse with rapidly decreasing prices [Sornette, 2004]).

If reflection is applied to a sequence $p = \ell + \nu$ that is the sum of a log-periodic component and a noise component, reflection will apply indiscriminately to both components. Thus, it is possible in principle that reflection would alter the power spectrum of noise. However, several Monte Carlo simulations showed that reflection leaves the power spectrum of Ornstein-Uhlenbeck noise practically unaffected. For example, Figure 5 gives the DFT amplitude of a realization of an Ornstein-Uhlenbeck process, the DFT amplitude of the reflection of the same process, and $\sqrt{S(f)}$. Reflection leaves the same general spectral behavior because, intuitively, the autocorrelation of stochastic signals is invariant to time reversal and the autocorrelation of an Ornstein-Uhlenbeck process decays exponentially.

5 Log-Periodic Spectra

To derive the LPPL spectrum, we will break down the LPPL model into its two main components of power law (without oscillations) and log-periodicity (without super-exponential growth). Then, we will return to the general case with the insight gained from the special cases, and verify numerically its spectral behavior.

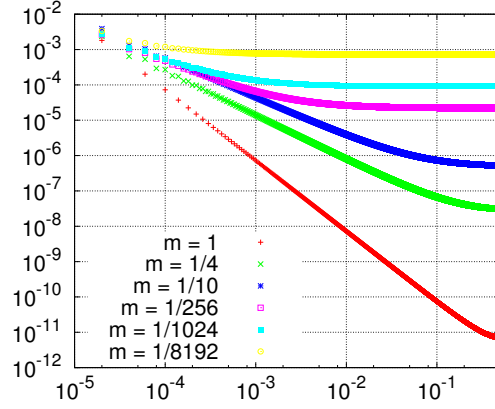


Figure 6: DFT amplitude of the power law ℓ_m .

5.1 Power Law

Consider first the case when $C = 0$, i.e., the LPPL function is completely determined by its power law behavior. If $m = C = 0$ (or if $B = 0$), ℓ reduces to a constant, and its transform $L(f) = 0$ for $f \neq 0$. If $C = 0, m = 1$, the reflected LPPL reduces to a triangular wave, for which it is known that $|L(f_i)| \propto 1/f_i^2$ when i is odd and $L(f_i) = 0$ when $i > 0$ is even.

The super-exponential behavior is the most pronounced when $m \rightarrow 0^+$, and A, B change as a function of m so that ℓ is not a constant. First, suppose for the time being that $T = n$ (the same result will hold for $T > n$ except for changes in the normalization constants). If A is multiplied by a factor independent of t , $L(f)$ remains unchanged ($f \neq 0$). Similarly, if B is multiplied by a factor independent of t , $L(f)$ is multiplied by the same factor ($f \neq 0$). In other words, a normalization of the constants A and B renormalizes $L(f)$ but does not alter its qualitative behavior. Assume that A and B are renormalized so that $\ell(0) = 0$ and $\int_0^T \ell(t) dt = 1/2$. Thus, $A = \hat{A}(m) = (1 + 1/m)/(2T)$ and $B = \hat{B}(m) = A/T^m$. Let $\ell_m(t) = \hat{A}(m) - \hat{B}(m)(T - t)^m$. Note that $\lim_{m \rightarrow 0^+} \hat{A}(m) = \infty$ and that we take the transform of reflected signals. Hence, $\lim_{m \rightarrow 0^+} \ell_m(t)$ is the delta function, and its spectrum is a constant (specifically, $1/2T$ with the common conventions on the DFT normalization factor). Figure 6 depicts the convergence of ℓ_m to the delta function in the frequency domain. When $m \neq 0$ is small, a flat spectrum characterizes the higher frequencies. As $m \rightarrow 0^+$, a constant spectrum is found in a progressively larger frequency interval. Meanwhile, the spectrum at low frequencies has an increasingly flatter slope: $|L(f)| \propto 1/f^2$ at $m = 1$, $|L(f)| \propto 1/f$ at $m \simeq 0.1$, and so on. The amplitude $|L(f)|$ decreases faster than $1/f$ for the range $m \geq 0.1$ that defines the “stylized features of LPPL” [Lin et al., 2009].

The same behavior holds also when $n < T$ by appropriately scaling the normalization constants of A and B . The same results hold qualitatively for A, B constant: the spectrum is flat although the absolute values do change depending on the normalization factor (see, for example, Figure 7).

5.2 Log-Periodicity

If $m = 0$, then $\ell(t) = A - B(1 + C \cos(\omega \ln(T - t) + \phi))$ (see Figure 8 for an example). By discounting the constant terms, the resulting signal is $\tilde{\ell}(t) = \cos(\omega \ln(T - t) + \phi)$. The signal $\tilde{\ell}(t)$ includes a

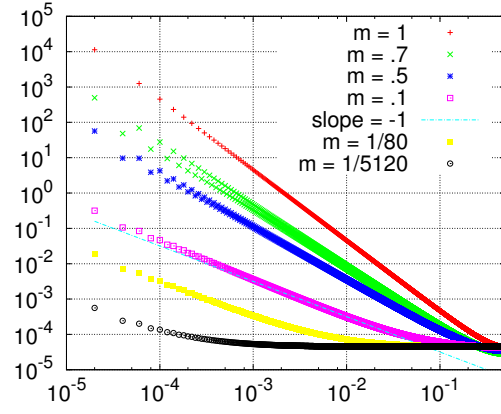


Figure 7: DFT amplitude of an LPPL function for various values of m ($A = 10, B = 0.01, T = n = 25000, C = 0$).

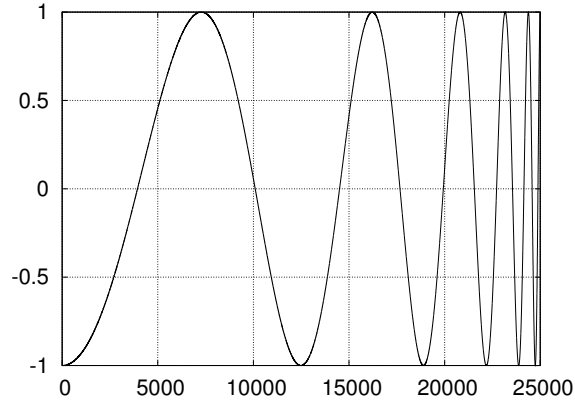


Figure 8: An LPPL function that highlights the oscillatory term as a function of non-linear time scaling ($A = 1, B = 1, C = -1, m = 0, \omega = 3\pi, n = 25000, T = n/(1 - e^{-11\pi/\omega}), \phi = -\omega \ln(T - n)$).

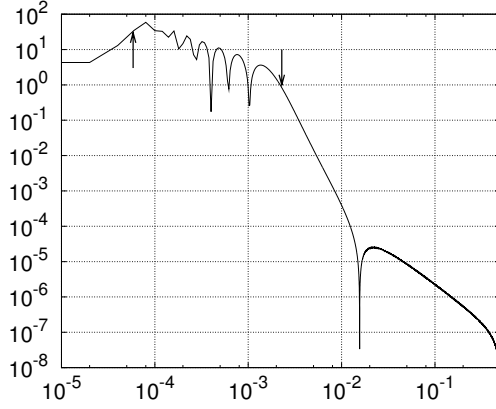


Figure 9: The DFT amplitude of the LPPL function in Figure 8, with arrows pointing to the minimum and maximum frequencies of the log-periodic spectrum.

non-linear time scaling, which can be interpreted as the frequency modulation of a carrier signal $\sin(2\pi f(t)t + \alpha)$ with some underlying signal $f(t)$. Given a frequency modulated signal, such as $\tilde{\ell}(t)$, its power spectrum resides almost entirely in the band between the maximum and minimum value of the modulating signal [Carson, 1922]. In practice, the differential frequency at time t is given by the derivative of phase $\omega \ln(T-t) + \phi$ normalized by 2π , and is thus equal to $\omega/(2\pi(T-t))$. Hence, the minimum frequency is $\omega/(2\pi T)$ and the maximum frequency is $\omega/(2\pi(T-n))$. Figure 9 shows an example along with the predicted frequency bounds.

Frequency-modulated signals are known to be resilient to changes in signal amplitude. In particular, the spectrum of $\tilde{\ell}_m(t) = (T-t)^m \cos(\omega \ln(T-t) + \phi)$ should be qualitatively close to the spectrum of $\tilde{\ell}_0(t)$. For example, Figure 10 shows the spectrum of Figure 9 except that now $m = 0.3$. The maximum and minimum frequencies still delimit the area where the frequency modulated signal has predominant power.

Incidentally, the critical time T is in a special position in that it affects both power law (due to $(T-t)^m$ term) and log-periodicity (due to the $\ln(T-t)$ term). Correspondingly, T determines both the maximum frequency $\omega/(2\pi(T-n))$ in the frequency modulated spectrum and the normalization constants $\hat{A}(m), \hat{B}(m)$ in the power law spectrum. In both cases, the effect of T is similar in that, when T approaches n , the maximum frequency and the normalization constants both increase.

Frequency modulation analysis could be regarded as the dual of the Lomb transform of the de-trended signal [Jiang et al., 2010] in that the frequency bounds are useful to estimate log-periodicity, whereas Lomb analysis is useful to confirm it after the LPPL parameters have already been estimated.

5.3 Combined LPPL

The two LPPL hallmarks are power law and log-periodicity. Each translates directly into the frequency domain. The power law implies a flat spectrum at high frequencies and moderate slope at low frequencies. Log-periodicity implies a bounded spectrum similar to that of a frequency modulated signal. Given the spectrum $|P|^2$ of prices during a bubble, we would expect to find either (or both) signatures of LPPL in the frequency domain.

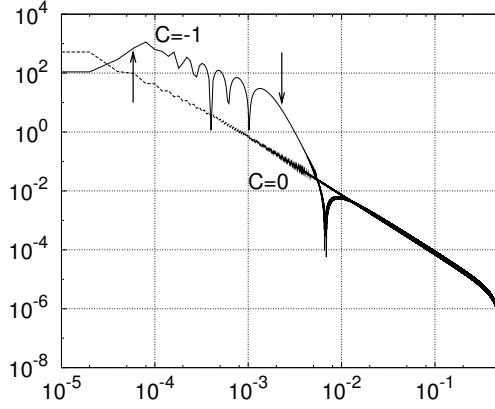


Figure 10: The DFT amplitude of the LPPL function in Figure 8 but with $m = 0.3$, with arrows pointing to the minimum and maximum frequencies of the log-periodic spectrum.

The general LPPL spectrum is analytically intractable, and it was estimated numerically. Sequences were generated for hundreds of different LPPL parameter sets, the respective spectra were calculated numerically and plotted. In general, to make sense of the spectrum, it is helpful to interpret it qualitatively in terms of the two main features of power law and log-periodicity. Specifically, a LPPL spectrum can be thought as the superposition of a power law spectrum and a log-periodic spectrum, where the superposition is weighted by the parameter C . For example, Figure 10 compares an LPPL spectrum with that of a pure power law with the same parameter values (but $C = 0$). As another example, Figure 11 shows the spectrum for a pure power law function ($m = 0.7, C = 0$), a pure log-periodic function ($m = 0, C = 0.5$), and two intermediate cases ($m = 0.7, C = 0.01, 0.05$). The intermediate cases have a spectrum that is visually the superposition of the two extreme spectra. Furthermore, higher values of C make the log-periodic component more visible. If C is small, the frequency modulation lobes are typically visible around the trend rather than above it (e.g., Figure 12). Results were similar for all other combinations of parameters, and omitted. Although it is helpful to comprehend a general LPPL spectrum as a superposition of feature, the intuition can seldom be translated into a closed expression. However, it is always useful to think the LPPL spectrum as the combination of two broad features. In particular, given a noisy LPPL, we would expect to find the two LPPL signatures in the frequency domain.

6 De-Noising

We now turn to discuss the extent to which it is possible to extract an underlying LPPL signal from noisy price measurements. The expectation would intuitively be that the mean-reverting noise smooths out relatively quickly so that the underlying LPPL behavior should be visible at the lower frequencies. More precisely, the potential for de-noising depends on achieving high values of the *signal-to-noise ratio* $R(f) = |L(f)|^2/S(f)$. The signal-to-noise ratio $R(f)$ is related to the non-causal *Weiner filter* $K(f) = 1/(1 + 1/R(f))$. The Wiener filter K has the property that the filtered signal $\hat{\ell}(i) = \sum_{j=-\infty}^{\infty} k(j)p(i-j)$ minimizes the mean square error $\sum (\ell(i) - \hat{\ell}(i))^2$ among all linear filters, where the k 's are the inverse transform of K and p is the reflected and periodicized

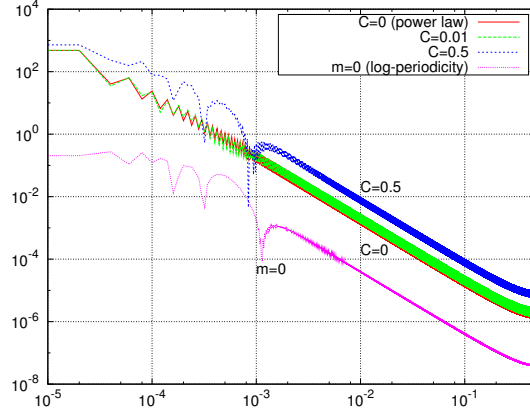


Figure 11: $|L(f)|$ for $A = 100, B = 0.01, T = 26000, n = 25000, \omega = 2\pi, \phi = 0$ in the cases of pure power law ($m = 0.7, C = 0$), pure log-periodicity ($m = 0, C = 0.5$), and two intermediate cases ($m = 0.7, C = 0.01, 0.5$).

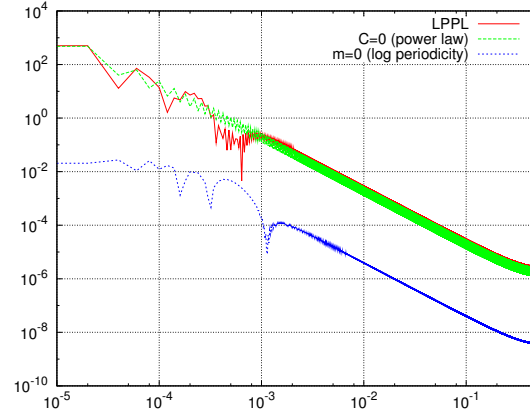


Figure 12: $|L(f)|$ for $A = 100, B = 0.01, T = 26000, n = 25000, \omega = 2\pi, \phi = 0$ in the cases of pure power law ($m = 0.7, C = 0$), pure log-periodicity ($m = 0, C = 0.05$), and combined case ($m = 0.7, C = 0.05$).

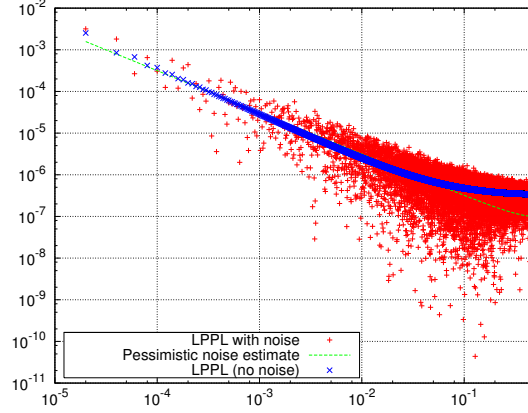


Figure 13: Spectrum $|L(f)|$ of $\ell_{0.1}$ with and without noise, and pessimistic noise estimate.

time series of prices. If $R(f)$ decreases with f , then $K(f)$ is a low pass filter and, intuitively, the filter replace each price $p(i)$ with a smoothed combination of its neighbors so as to reject the noise. The signal-to-noise ratio is a function of frequency and can take various shapes depending on the parameters of the underlying LPPL ℓ and Ornstein-Uhlenbeck noise.

In the stylized LPPL, $m \geq 0.1$, and, with a pure power law ($C = 0$), $|L|^2$ always decreases faster than the slope of S . Hence, the power law contributes to de-noising only if $|L(f)|^2 > S(f)$ at low frequencies. In pure log-periodicity ($m = 0$), $|L(f)|$ is significant only for frequencies $f \leq \omega/(2\pi(T - n))$. Again, log-periodicity contributes to de-noising only if $|L(f)|^2 > S(f)$ at low frequencies.

It is useful to examine the special case when $|L(f)| \propto 1/f$ at low frequencies since $|L|$'s slope matches exactly the slope of S_∞ . In this case, it may be hard to ascertain the presence of a signal ℓ if Wiener noise is also present since the LPPL spectrum almost perfectly overlaps with the noise spectrum. Figure 13 shows the amplitude $|L|$ of an LPPL $\ell_{0.1}$ without noise. Then, Gaussian white noise was generated with variance σ^2 chosen so that the corresponding spectrum S_∞ matches closely $|L|^2$ at low frequencies. The amplitude of the resulting price transform is shown in Figure 13 along with the pessimistic estimate defined in Section 3.4. The pessimistic estimate overlapped almost exactly the true noise spectrum (whereas the maximum likelihood estimate did not produce finite parameter estimates). The spectra $|L|^2$, $|P|^2$, and S_∞ were practically identical at low frequencies, but at high frequencies $|P|$ followed the larger $|L|$. We conclude that even in the special case when the LPPL is hidden under the noise because m makes $|L(f)| \propto 1/f$ and σ^2 makes $|L(f)|^2 \simeq S_\infty$ at low frequencies, LPPL and noise can be distinguished by extracting a pessimistic noise estimate. However, in this case, the difference is noticeable at high frequencies.

In summary, the intuition was that LPPL should be visible at the low frequencies, and it is correct in most cases. Then, de-noising should be possible when the spectrum looks qualitatively as in Figure 4, with the price spectrum clearly above the noise estimate for a significant range of frequencies at the low end of the spectrum. However, we found special combinations of parameters (e.g., Figure 13) where the presence of LPPL is visible only from the highest frequencies.

If the time horizon n approaches the critical time T , intuition would suggest that it should be easier to detect an LPPL trend. Indeed, if the critical time T is approaching, then the max-

Series	Period	Data Points
Dow Jones Industrial Average	June 1921-July 1929	2440
S&P 500	July 1985-July 1987	527
NASDAQ Composite	January 1994-February 2000	1555
S&P 500	July 2003-June 2007	1000
GLD	March 2009-October 2009	171

Table 1: Price time series used in the evaluation.

imum log-periodic frequency is higher resulting into a wider frequency modulated spectrum, and the normalization constant \hat{B} is larger, shifting the spectrum upward. In other words, when the critical time approaches, the LPPL function resembles more closely an impulse and additionally its oscillations are wilder, so that it is hard for mean-reverting noise to disguise the underlying LPPL trend.

7 Evaluation

Data sets are daily closing values of the major indexes and securities shown in Table 1. The first four data sets are associated with major historical bubbles in the U.S. stock market. The last data set is a recent bubble in gold prices that burst in November 2009 [Sornette et al., 2009]. These prices sequences stop closely before the time at which the price reached its maximum level during the bubble episode. Therefore, the detection of an underlying LPPL trend should be relatively easy. We have additionally considered individual stocks. For example, we have investigated various tech stocks during the 1998-2000 bubble and the Netflix stock in the 2005-10 period. The results are qualitatively similar to those reported here, and omitted.

The maximum likelihood and pessimistic noise estimates were close to each other, providing evidence that mean reversion is weak. Both estimates fit well the price spectrum, a sign that prices can be explained primarily as a Wiener process. In most cases, the spectrum deviates from the pessimistic estimate only at $f = 0, 1/(2n)$, which is consistent with adding Wiener noise to deterministic exponential price trajectory.

As for the maximum likelihood estimate, the diffusion constant is relatively large. The maximum likelihood estimate was slightly but consistently smaller at low frequencies than the price spectrum in two cases: Dow Jones 1929 and gold 2009. The discrepancy was investigated with a filter to remove the maximum likelihood noise. The filter is non-causal and produces a signal $\tilde{\ell}$ with the property that $\tilde{L}(f) = L(f)$ if $f < \tilde{f}$ and $\tilde{L}(f) = 0$ otherwise, where \tilde{f} is the smallest frequency f at which $|L(f)|^2 < S(f)$. The intuition is that the underlying LPPL spectrum $|L|^2$ decreases faster than the noise spectrum S , so if one takes the smallest frequency f at which $|L(f)|^2 \geq S(f)$, the resulting low-pass filter approximates a Wiener filter. Figures 19 and 20 show the original and filtered series. Both filtered series show significant oscillations at the beginning and slight super-exponential growth throughout. However, the oscillations are absent toward the end of the filtered sequence. Since log-periodic oscillations become progressively more rapid toward the critical time, even if log-periodicity were present in these price sequences, oscillations would be blocked by the filter at the highest log-periodic frequencies. However, the filter only blocks frequencies when noise dominate the signal, and so we conclude that log-periodicity, if at all present, has been disrupted by

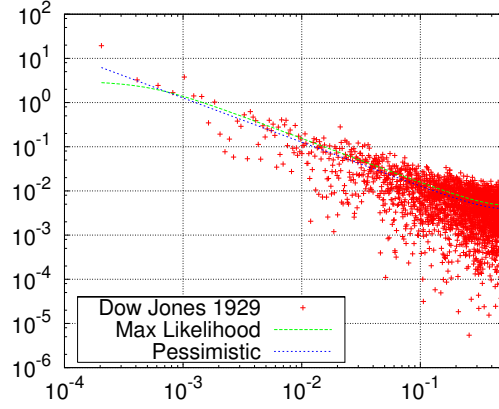


Figure 14: Spectrum of the Dow Jones 1929 price sequence.

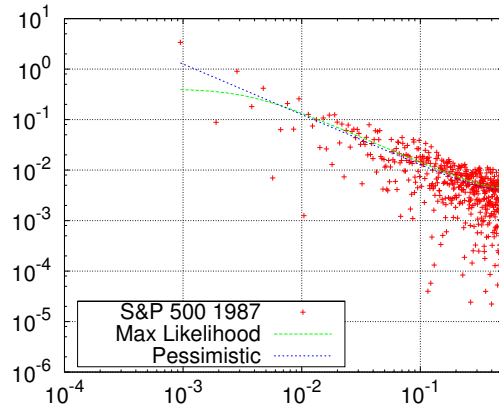


Figure 15: Spectrum of the S&P 500 1987 price sequence.

noise. Similarly, super-exponential growth is visually unclear both in the time and in the frequency domain.

In summary, noise had at best weak mean reversion and relatively high variance, so that, even if LPPL underlied price dynamics, it was obscured by noise.

8 Conclusions

The estimators show that mean reversion is weak, and Ornstein-Uhlenbeck noise is close to a non-mean-reverting Wiener process. Furthermore, the signal-to-noise ratio $R(f)$ is low, and it mostly makes it possible to reconstruct an underlying exponential trend. Even in those cases where $R(f) > 0$ for a range of low frequencies, noise makes it impossible to reconstruct either log-periodicity or a power law. In short, due to noise, there is no trace of LPPL during bubbles.

However, it is possible in principle that LPPL underlies price dynamics, but that it cannot be isolated with pure frequency methods. Future work will investigate different transforms, such

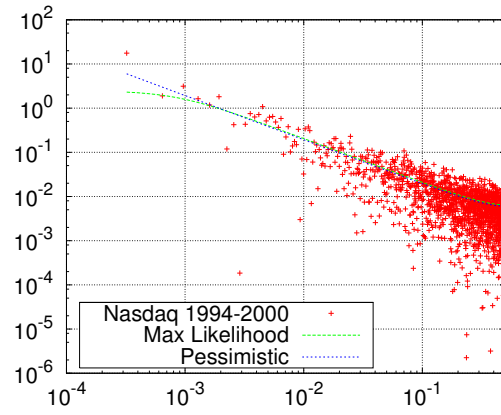


Figure 16: Spectrum of the NASDAQ Composite 1994-2000 price sequence.

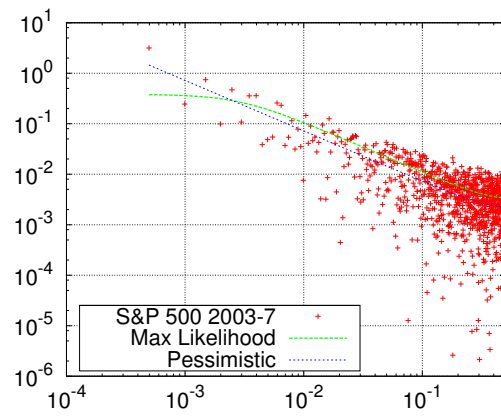


Figure 17: Spectrum of the S&P 500 2003-07 price sequence.

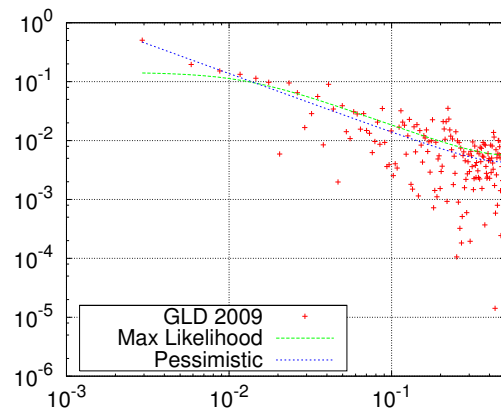


Figure 18: Spectrum of the GLD (gold) 2009 price sequence.

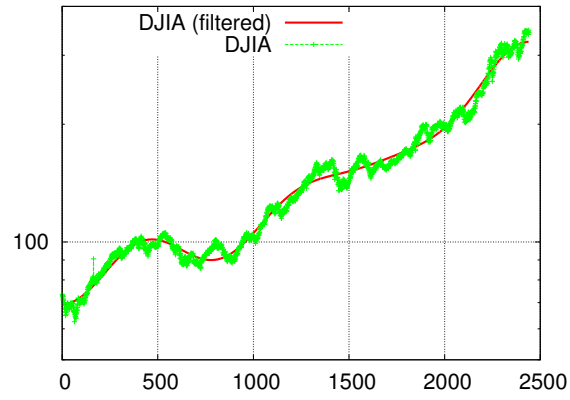


Figure 19: Dow Jones 1929 with filtered price sequence ($\tilde{f} = f_9$).

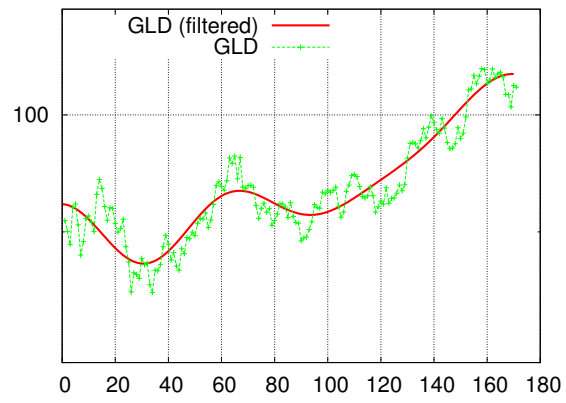


Figure 20: Gold 2009 with filtered price sequence ($\tilde{f} = f_6$).

as wavelets, that may make it possible to discern explosive bubble growth. Furthermore, if the underlying LPPL were expressed as a dynamic system with known parameters, it would be possible to use the additional information in de-noising filters that extend a pure frequency approach.

References

- [Bibbona et al., 2008] Bibbona, E., Panfilo, G., and Taveggia, P. (2008). The Ornstein-Uhlenbeck process as a model of a low pass filtered white noise. *Metrologia*, 45:S117–S126.
- [Brèe et al., 2010] Brèe, D., Challet, D., and Peirano, P. (2010). Prediction accuracy and sloppiness of log-periodic functions. Technical Report 1006:2010v1, arXiv.
- [Carson, 1922] Carson, J. R. (1922). Notes on the theory of modulation. *Proc. IRE*, 10(1):57–64.
- [Chang and Feigenbaum, 2006] Chang, G. and Feigenbaum, J. (2006). A Bayesian analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 6(1):15–36.
- [Chang and Feigenbaum, 2008] Chang, G. and Feigenbaum, J. (2008). Detecting log-periodicity in a regime-switching model of stock returns. *Quantitative Finance*, 8:723–738.
- [Feigenbaum, 2001] Feigenbaum, J. A. (2001). A statistical analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 1:346–360.
- [Gillespie, 1991] Gillespie, D. (1991). *Markov processes: An introduction for physical scientists*. Academic Pr.
- [Granger and Newbold, 1974] Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2:111–120.
- [Jiang et al., 2010] Jiang, Z.-Q., Zhou, W.-X., Sornette, D., Woodard, R., Bastiaensen, K., and Cauwels, P. (2010). Bubble diagnosis and prediction of the 2005-2007 and 2008-2009 Chinese stock market bubbles. *Economic Behavior and Organization*, 74:149–162.
- [Liberatore, 2010] Liberatore, V. (2010). Computational LPPL fit to financial bubbles. Technical Report 1003.2920, ArXiv.
- [Lin et al., 2009] Lin, L., E, R. R., and Sornette, D. (2009). A Consistent Model of ‘Explosive’ Financial Bubbles With Mean-Reversing Residuals. *ArXiv e-prints*.
- [Loève, 1978] Loève, M. (1978). *Probability theory II*, volume 46 of *Graduate Texts in Mathematics*. Springer-Verlag, 4 edition.
- [Phillips, 1986] Phillips, P. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, 31:311–340.
- [Pollock, 2001] Pollock, D. S. G. (2001). Methodology for trend estimation. *Economic Modelling*, 18(1):75 – 96.
- [Råde and Westergren, 1999] Råde, L. and Westergren, B. (1999). *Mathematics Handbook for Science and Engineering*. Springer, Berlin.

- [Sornette, 2004] Sornette, D. (2004). *Critical phenomena in natural sciences: chaos, fractals, self-organization, and disorder: concepts and tools*. Springer.
- [Sornette et al., 2009] Sornette, D., Woodard, R., Fedorovsky, M., Reimann, S., Woodard, H., and Zhou, W.-X. (2009). The financial bubble experiment: advanced diagnostics and forecasts of bubble terminations.
- [Stauffer and Sornette, 1998] Stauffer, D. and Sornette, D. (1998). Log-periodic oscillations for biased diffusion on random lattice. *Physica A*, 252:271–277.
- [van den Berg, 2007] van den Berg, M. (2007). Calibrating the Ornstein-Uhlenbeck model.
- [Wang and Uhlenbeck, 1945] Wang, M. C. and Uhlenbeck, G. E. (1945). On the theory of Brownian motion II. *Rev. Mod. Phys.*, 17:323–342.